

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.**

Application Serial Number: 10/665,184A  
Source: 1Fw/6  
Date Processed by STIC: 3/2/06

# ***ENTERED***



IFW16

## RAW SEQUENCE LISTING

DATE: 03/02/2006

PATENT APPLICATION: US/10/665,184A

TIME: 11:32:30

Input Set : A:\24348-501CIP.ST25.txt

Output Set: N:\CRF4\03022006\J665184A.raw

3 <110> APPLICANT: Ben-Sasson, Shmuel  
 4 Cohen, Einat  
 6 <120> TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration  
 Across a  
 7 Biological Barrier  
 9 <130> FILE REFERENCE: 24348-501CIP  
 11 <140> CURRENT APPLICATION NUMBER: 10/665,184A  
 12 <141> CURRENT FILING DATE: 2003-09-17  
 14 <150> PRIOR APPLICATION NUMBER: PCT/IB03/00968  
 15 <151> PRIOR FILING DATE: 2003-02-07  
 17 <150> PRIOR APPLICATION NUMBER: 60/355,396  
 18 <151> PRIOR FILING DATE: 2002-02-07  
 20 <160> NUMBER OF SEQ ID NOS: 66  
 22 <170> SOFTWARE: PatentIn version 3.2  
 24 <210> SEQ ID NO: 1  
 25 <211> LENGTH: 23  
 26 <212> TYPE: PRT  
 27 <213> ORGANISM: haemophilus influenzae  
 29 <400> SEQUENCE: 1  
 31 Asn Tyr His Asp Ile Val Leu Ala Leu Ala Gly Val Cys Gln Ser Ala  
 32 1 5 10 15  
 35 Lys Leu Val His Gln Leu Ala  
 36 20  
 39 <210> SEQ ID NO: 2  
 40 <211> LENGTH: 23  
 41 <212> TYPE: PRT  
 42 <213> ORGANISM: Pasteurella multocida  
 44 <400> SEQUENCE: 2  
 46 Asn Tyr Tyr Asp Ile Thr Leu Ala Leu Ala Gly Val Cys Gln Ala Ala  
 47 1 5 10 15  
 50 Lys Leu Val Gln Gln Phe Ala  
 51 20  
 54 <210> SEQ ID NO: 3  
 55 <211> LENGTH: 23  
 56 <212> TYPE: PRT  
 57 <213> ORGANISM: Escherichia coli  
 59 <400> SEQUENCE: 3  
 61 Asn Tyr Tyr Asp Ile Thr Leu Ala Leu Ala Gly Ile Cys Gln Ser Ala  
 62 1 5 10 15  
 65 Arg Leu Val Gln Gln Leu Ala  
 66 20  
 69 <210> SEQ ID NO: 4  
 70 <211> LENGTH: 23  
 71 <212> TYPE: PRT

p.6

## RAW SEQUENCE LISTING

DATE: 03/02/2006

PATENT APPLICATION: US/10/665,184A

TIME: 11:32:30

Input Set : A:\24348-501CIP.ST25.txt

Output Set: N:\CRF4\03022006\J665184A.raw

```

72 <213> ORGANISM: Vibrio cholerae
74 <400> SEQUENCE: 4
76 Ala Ile Tyr Asp Arg Thr Ile Ala Phe Ala Gly Ile Cys Gln Ala Val
77 1           5           10           15
80 Ala Leu Val Gln Gln Val Ala
81           20
84 <210> SEQ ID NO: 5
85 <211> LENGTH: 23
86 <212> TYPE: PRT
87 <213> ORGANISM: Buchnera aphidicola
89 <400> SEQUENCE: 5
91 Lys Ile His Leu Ile Thr Leu Ser Leu Ala Gly Ile Cys Gln Ser Ala
92 1           5           10           15
95 His Leu Val Gln Gln Leu Ala
96           20
99 <210> SEQ ID NO: 6
100 <211> LENGTH: 23
101 <212> TYPE: PRT
102 <213> ORGANISM: Pseudomonas aeruginosa
104 <400> SEQUENCE: 6
106 Asp Pro Arg Gln Gln Leu Ile Ala Leu Gly Ala Val Phe Glu Ser Ala
107 1           5           10           15
110 Ala Leu Val Asp Lys Leu Ala
111           20
114 <210> SEQ ID NO: 7
115 <211> LENGTH: 23
116 <212> TYPE: PRT
117 <213> ORGANISM: Xylella fastidiosa
119 <400> SEQUENCE: 7
121 Leu Ile Asp Asn Arg Val Leu Ala Leu Ala Gly Val Val Gln Ala Leu
122 1           5           10           15
125 Gln Gln Val Arg Gln Ile Ala
126           20
129 <210> SEQ ID NO: 8
130 <211> LENGTH: 23
131 <212> TYPE: PRT
132 <213> ORGANISM: Rhizobium loti
134 <400> SEQUENCE: 8
136 Asn Leu Pro Pro Ile Val Leu Ala Val Ile Gly Ile Cys Ala Ala Val
137 1           5           10           15
140 Phe Leu Leu Gln Gln Tyr Val
141           20
144 <210> SEQ ID NO: 9
145 <211> LENGTH: 23
146 <212> TYPE: PRT
147 <213> ORGANISM: Homo sapiens
149 <400> SEQUENCE: 9
151 Asn Tyr Phe Ile Val Asn Leu Ala Leu Ala Asp Leu Cys Met Ala Ala
152 1           5           10           15

```

## RAW SEQUENCE LISTING

DATE: 03/02/2006

PATENT APPLICATION: US/10/665,184A

TIME: 11:32:30

Input Set : A:\24348-501CIP.ST25.txt

Output Set: N:\CRF4\03022006\J665184A.raw

```

155 Phe Asn Ala Ala Phe Asn Phe
156      20
159 <210> SEQ ID NO: 10
160 <211> LENGTH: 23
161 <212> TYPE: PRT
162 <213> ORGANISM: Chlamydia pneumoniae
164 <400> SEQUENCE: 10
166 Thr Ala Phe Asp Phe Asn Lys Met Leu Asp Gly Val Cys Thr Tyr Val
167 1      5      10      15
170 Lys Gly Val Gln Gln Tyr Leu
171      20
174 <210> SEQ ID NO: 11
175 <211> LENGTH: 23
176 <212> TYPE: PRT
177 <213> ORGANISM: Rhizobium loti
179 <400> SEQUENCE: 11
181 Arg Ala Ile Leu Ile Pro Leu Ala Leu Ala Gly Leu Cys Gln Val Ala
182 1      5      10      15
185 Arg Ala Gly Asp Ile Ser Ser
186      20
189 <210> SEQ ID NO: 12
190 <211> LENGTH: 25
191 <212> TYPE: PRT
192 <213> ORGANISM: Bacillus subtilis
194 <400> SEQUENCE: 12
196 Met Arg Asn Leu Thr Lys Thr Ser Leu Leu Leu Ala Gly Leu Cys Thr
197 1      5      10      15
200 Ala Ala Gln Met Val Phe Val Thr His
201      20      25
204 <210> SEQ ID NO: 13
205 <211> LENGTH: 25
206 <212> TYPE: PRT
207 <213> ORGANISM: Kingella denitrificans
209 <400> SEQUENCE: 13
211 Ile Glu Leu Met Ile Val Ile Ala Ile Ile Gly Ile Leu Ala Ala Ile
212 1      5      10      15
215 Ala Leu Pro Ala Tyr Gln Glu Tyr Val
216      20      25
219 <210> SEQ ID NO: 14
220 <211> LENGTH: 25
221 <212> TYPE: PRT
222 <213> ORGANISM: Eikenella corrodens
224 <400> SEQUENCE: 14
226 Ile Glu Leu Met Ile Val Ile Ala Ile Ile Gly Ile Leu Ala Ala Ile
227 1      5      10      15
230 Ala Leu Pro Ala Tyr Gln Asp Tyr Val
231      20      25
234 <210> SEQ ID NO: 15
235 <211> LENGTH: 16

```

## RAW SEQUENCE LISTING

DATE: 03/02/2006

PATENT APPLICATION: US/10/665,184A

TIME: 11:32:30

Input Set : A:\24348-501CIP.ST25.txt

Output Set: N:\CRF4\03022006\J665184A.raw

236 <212> TYPE: PRT  
 237 <213> ORGANISM: zonula occludens toxin  
 239 <400> SEQUENCE: 15  
 241 Ala Ser Phe Gly Phe Cys Ile Gly Arg Leu Cys Val Gln Asp Gly Phe  
 242 1 5 10 15  
 245 <210> SEQ ID NO: 16  
 246 <211> LENGTH: 6  
 247 <212> TYPE: PRT  
 248 <213> ORGANISM: Artificial  
 250 <220> FEATURE:  
 251 <223> OTHER INFORMATION: Linker  
 253 <400> SEQUENCE: 16  
 255 Gly Gly Lys Gly Gly Lys  
 256 1 5  
 259 <210> SEQ ID NO: 17  
 260 <211> LENGTH: 5  
 261 <212> TYPE: PRT  
 262 <213> ORGANISM: Artificial  
 264 <220> FEATURE:  
 265 <223> OTHER INFORMATION: Fully conserved "strong" amino acid residue chain  
 267 <400> SEQUENCE: 17  
 269 Asn Arg Glu Gln Lys  
 270 1 5  
 273 <210> SEQ ID NO: 18  
 274 <211> LENGTH: 4  
 275 <212> TYPE: PRT  
 276 <213> ORGANISM: Artificial  
 278 <220> FEATURE:  
 279 <223> OTHER INFORMATION: Fully conserved "strong" amino acid residue chain  
 281 <400> SEQUENCE: 18  
 283 Asn His Gln Lys  
 284 1  
 287 <210> SEQ ID NO: 19  
 288 <211> LENGTH: 4  
 289 <212> TYPE: PRT  
 290 <213> ORGANISM: Artificial  
 292 <220> FEATURE:  
 293 <223> OTHER INFORMATION: Fully conserved "strong" amino acid residue chain  
 295 <400> SEQUENCE: 19  
 297 Asn Asp Glu Gln  
 298 1  
 301 <210> SEQ ID NO: 20  
 302 <211> LENGTH: 4  
 303 <212> TYPE: PRT  
 304 <213> ORGANISM: Artificial  
 306 <220> FEATURE:  
 307 <223> OTHER INFORMATION: Fully conserved "strong" amino acid residue chain  
 309 <400> SEQUENCE: 20  
 311 Gln His Arg Lys

## RAW SEQUENCE LISTING

DATE: 03/02/2006

PATENT APPLICATION: US/10/665,184A

TIME: 11:32:30

Input Set : A:\24348-501CIP.ST25.txt

Output Set: N:\CRF4\03022006\J665184A.raw

```

312 1
315 <210> SEQ ID NO: 21
316 <211> LENGTH: 4
317 <212> TYPE: PRT
318 <213> ORGANISM: Artificial
320 <220> FEATURE:
321 <223> OTHER INFORMATION: Fully conserved "strong" amino acid residue chain
323 <400> SEQUENCE: 21
325 Met Ile Leu Val
326 1
329 <210> SEQ ID NO: 22
330 <211> LENGTH: 30
331 <212> TYPE: PRT
332 <213> ORGANISM: Artificial
334 <220> FEATURE:
335 <223> OTHER INFORMATION: Penetrating peptide
338 <220> FEATURE:
339 <221> NAME/KEY: MOD_RES
340 <222> LOCATION: (1)..(1)
341 <223> OTHER INFORMATION: ACETYLATION
343 <220> FEATURE:
344 <221> NAME/KEY: PEPTIDE
345 <222> LOCATION: (30)..(30)
346 <223> OTHER INFORMATION: wherein Xaa is Lysine-NH2
348 <220> FEATURE:
349 <221> NAME/KEY: misc_feature
350 <222> LOCATION: (30)..(30)
351 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
353 <400> SEQUENCE: 22
355 Asn Tyr Tyr Asp Ile Thr Leu Ala Leu Ala Gly Ile Cys Gln Ser Ala
356 1          5          10          15
W--> 359 Arg Leu Val Gln Gln Leu Ala Gly Gly Gly Lys Gly Gly Xaa
360          20          25          30
363 <210> SEQ ID NO: 23
364 <211> LENGTH: 4
365 <212> TYPE: PRT
366 <213> ORGANISM: Artificial
368 <220> FEATURE:
369 <223> OTHER INFORMATION: Fully conserved "strong" amino acid residue chain
371 <400> SEQUENCE: 23
373 Met Ile Leu Phe
374 1
377 <210> SEQ ID NO: 24
378 <211> LENGTH: 23
379 <212> TYPE: PRT
380 <213> ORGANISM: Homo sapiens
382 <400> SEQUENCE: 24
384 Asn Tyr Phe Leu Val Asn Leu Ala Phe Ala Glu Ala Ser Met Ala Ala
385 1          5          10          15

```

RAW SEQUENCE LISTING ERROR SUMMARY      DATE: 03/02/2006  
 PATENT APPLICATION: US/10/665,184A      TIME: 11:32:31

Input Set : A:\24348-501CIP.ST25.txt  
 Output Set: N:\CRF4\03022006\J665184A.raw

**Please Note:**

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:22; Xaa Pos. 30  
 Seq#:30; Xaa Pos. 30  
 Seq#:31; Xaa Pos. 30  
 Seq#:32; Xaa Pos. 30  
 Seq#:33; Xaa Pos. 29  
 Seq#:34; Xaa Pos. 29  
 Seq#:35; Xaa Pos. 31  
 Seq#:36; Xaa Pos. 30  
 Seq#:37; Xaa Pos. 29  
 Seq#:44; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16  
 Seq#:45; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
 Seq#:45; Xaa Pos. 23  
 Seq#:46; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
 Seq#:46; Xaa Pos. 23  
 Seq#:47; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
 Seq#:47; Xaa Pos. 23  
 Seq#:48; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
 Seq#:48; Xaa Pos. 23,24,25  
 Seq#:49; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
 Seq#:49; Xaa Pos. 23  
 Seq#:50; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
 Seq#:50; Xaa Pos. 23  
 Seq#:51; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
 Seq#:51; Xaa Pos. 23  
 Seq#:52; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
 Seq#:52; Xaa Pos. 23,24  
 Seq#:53; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
 Seq#:53; Xaa Pos. 23,24,25  
 Seq#:54; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
 Seq#:54; Xaa Pos. 23  
 Seq#:55; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
 Seq#:55; Xaa Pos. 23  
 Seq#:56; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
 Seq#:56; Xaa Pos. 23,24,25,26  
 Seq#:57; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
 Seq#:57; Xaa Pos. 23,24,25,26,27,28  
 Seq#:58; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22  
 Seq#:58; Xaa Pos. 23

**Invalid <213> Response:**

Use of "Artificial" only as "<213> Organism" response is incomplete,  
 per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:16,17,18,19,20,21,22,23,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45  
 Seq#:46,47,48,49,50,51,52,53,54,55,56,57,58,66

## VERIFICATION SUMMARY

DATE: 03/02/2006

PATENT APPLICATION: US/10/665,184A

TIME: 11:32:31

Input Set : A:\24348-501CIP.ST25.txt

Output Set: N:\CRF4\03022006\J665184A.raw

L:359 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22 after pos.:16  
L:497 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30 after pos.:16  
L:531 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31 after pos.:16  
L:565 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:32 after pos.:16  
L:599 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:33 after pos.:16  
L:633 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:34 after pos.:16  
L:667 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:35 after pos.:16  
L:701 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:36 after pos.:16  
L:735 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:37 after pos.:16  
L:919 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:44 after pos.:0  
L:1019 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:45 after pos.:0  
M:341 Repeated in SeqNo=45  
L:1103 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46 after pos.:0  
M:341 Repeated in SeqNo=46  
L:1187 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:47 after pos.:0  
M:341 Repeated in SeqNo=47  
L:1256 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:48 after pos.:0  
M:341 Repeated in SeqNo=48  
L:1350 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49 after pos.:0  
M:341 Repeated in SeqNo=49  
L:1414 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:50 after pos.:0  
M:341 Repeated in SeqNo=50  
L:1503 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:51 after pos.:0  
M:341 Repeated in SeqNo=51  
L:1617 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:52 after pos.:0  
M:341 Repeated in SeqNo=52  
L:1716 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:53 after pos.:0  
M:341 Repeated in SeqNo=53  
L:1810 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:54 after pos.:0  
M:341 Repeated in SeqNo=54  
L:1889 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:55 after pos.:0  
M:341 Repeated in SeqNo=55  
L:2003 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:56 after pos.:0  
M:341 Repeated in SeqNo=56  
L:2117 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:57 after pos.:0  
M:341 Repeated in SeqNo=57  
L:2196 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:58 after pos.:0  
M:341 Repeated in SeqNo=58